Методы пороговой криптографии для защиты облачных вычислений[1]

Варновский Н.П., Мартишин С.А., Храпченко М.В., Шокуров А.В.

Аннотация. Защита информации в облачных вычислениях активно исследуется мировым научным сообществом. Эти исследования показали, что обозначенная проблема намного сложнее тех задач защиты информации, которые решаются известными криптографическими средствами. Так, например в работе [1] рассмотрена математическая модель организации облачных вычислений и доказано, что уже в случае двух пользователей защита информации невозможна. Предлагается альтернативная модель организации облачных вычислений, в которой указанный отрицательный результат не имеет места. Исследуются методы защиты информации в этой новой модели.

Ключевые слова: защита информации, облачные вычисления, вычисления над зашифрованными данными, гомоморфные вычисления.

Проблемы защиты информации в облачных вычислениях стали активно исследоваться слишком поздно, когда эти вычисления уже стали фактически реализованной технологией. Только практические применения рассеяли ранее бытовавшие иллюзии, что в этом случае для защиты информации достаточно уже имеющихся криптографических средств.

Сложность возникающих проблем можно понять уже на простейших примерах. Обратимся к следующей модели. Облако рассматривается как единая сущность. Обозначим ее через S. В системе имеются также пользователи и клиенты. Для простоты будем считать, что множество пользователей $P_1,\ldots,P_n$ неизменно. Количество клиентов не регламентируется. У пользователя P_i имеются конфиденциальные данные x_i, хранящиеся на облаке (такой метод использования облака в литературе называют «базы данных как сервис»). В число пользователей может входить также само облако с данными x_S. Любой клиент C может обратиться к облаку с запросом на вычисление значения некоторой функции f, зависящей от конфиденциальных данных. Некоторые пользователи облака могут быть его клиентами. Множества пользователей и клиентов могут пересекаться. Запрос

[1] Работа поддержана грантом РФФИ 12-07-00206-а.

состоит из описания функции f, идентификатора клиента и его открытого ключа k_C. Облако должно проверить полномочия клиента C на вычисление $f(x_S, x_1, \ldots, x_n)$. Такая проверка реализуется с помощью стандартных криптографических средств и в данной статье не обсуждается. Если клиент C имеет право вычислять функцию f, то облако должно вычислить значение $E(k_C, f(x_S, x_1, \ldots, x_n))$ и отправить его клиенту. Здесь E - функция шифрования какой-либо криптосистемы с открытым ключом, например, RSA.

Описанная система может рассматриваться и как модель вычислений над конфиденциальными данными, и как модель базы данных с конфиденциальной информацией. В последнем случае функции f представляют собой запросы к базе данных.

Пользователь, который помещает в базу данных свои конфиденциальные данные, не доверяет облаку и должен обеспечить криптографическую защиту данных. Криптография подсказывает, казалось бы, надежное решение: данные следует зашифровать. Но здесь возникает следующая проблема. Если на облаке хранятся не данные x_i, а их криптограммы, то как вычислить функцию f?

В литературе можно найти описания подходов к организации вычислений над конфиденциальными данными. Но при этом совершенно упускается важнейший научный вопрос: а такая защита информации возможна?

Предполагаем, что противник имеет возможность обращаться к базе данных с конфиденциальной информацией с запросами двух типов:

- запросы на добавление в базу нового элемента;
- запросы на сравнение значений двух элементов.

Если интересующий противника элемент данных является целым числом из известного диапазона, то хорошо известным методом деления пополам значение этого элемента будет определено вне зависимости от используемой системы защиты информации.

Подобную угрозу можно пытаться предотвратить, вводя запреты на некоторые типы запросов. Мы далее рассмотрим предельный случай статичной базы данных – значения $x_1, \ldots, x_n$, хранящиеся на облаке, не меняются. Таким образом, речь идет именно о вычислениях над конфиденциальными данными.

В статье [4] для описанной выше модели облачных вычислений дано формальное определение стойкости системы защиты информации и доказано, что уже в случае двух пользователей стойкая защита конфиденциальных данных невозможна.

Этот результат опровергает прочно укоренившееся представление, что недавно предложенная [2] криптосистема гомоморфного шифрования решает, по крайней мере, теоретически все проблемы защиты информации в облачных вычислениях.

Мы предлагаем исследовать альтернативную модель облачных вычислений. Она отличается от вышеописанной тем, что каждый субъект, заинтересованный в обеспечении конфиденциальности, создает свой криптосервер. Обозначим эти криптосерверы $S_1,\ldots,S_m$. С точки зрения пользователей криптосерверы являются частью облака.

Для защиты данных, хранящихся на облаке, используется пороговая гомоморфная криптосистема с открытым ключом. Для ее инициализации криптосерверы выполняют специальный протокол, который создает пару ключей (k_1,k_2). Ключ k_1 - открытый: он доступен для всех пользователей. Ключ k_2 - секретный, он неизвестен никому. В результате выполнения протокола сервер S_i получает долю k_2^i секретного ключа k_2. Параметром криптосистемы является число t, $t<m$, которое называется порогом. Смысл этого параметра таков. Пусть E^t и D^t функции соответственно шифрования и дешифрования криптосистемы, и пусть $E^t(k_1,m)$ - криптограмма элемента данных m. Тогда любое подмножество из не менее чем t серверов может, используя свои доли секретного ключа, вычислить функцию D^t, то есть извлечь данные из криптограммы. А любая коалиция из не более, чем $(t-1)$ сервера, сложив свои доли k_2^i секретного ключа, не сможет этого сделать. Другими словами, криптосистема остается стойкой в присутствии противника, который контролирует не более $(t-1)$ серверов.

Необходимо подчеркнуть, что в процессе выполнения протокола, вычисляющего функцию D^t, секретный ключ k_2 никогда не восстанавливается из его долей. Все данные пользователей $x_1,\ldots,x_n$ шифруются все на одном ключе k_1. На облаке хранятся криптограммы $E^t(k_1,x_i)$.

Стойкость предлагаемой системы защиты информации основывается на стойкости пороговой гомоморфной криптосистемы и на следующем предположении: никакое подмножество из не менее чем t серверов не может вступить в сговор с целью нарушения конфиденциальности данных пользователей.

Благодаря этому предположению отрицательный результат [1] на нашу модель не распространяются.

Запрос клиента C выполняется следующим образом. Сначала облако, используя гомоморфность криптосистемы, вычисляет значение $E^t(k_1, E(k_C, f(x_S, x_1, \ldots, x_n)))$. Затем серверы выполняют протокол дешифрования, извлекая из этой криптограммы значение $E(k_C, f(x_S, x_1, \ldots, x_n))$.

Полное и математически строгое изложение будет приведено в отдельной статье.

Литература

[1] C. Gentry, Fully homomorphic encryption using ideal lattices , in Proceedings of the 41st ACM Symposium on Theory of Computing|STOC 2009, ACM, New York (2009), 169-178.

[2] C. Gentry and S. Halevi, Implementing Gentry's Fully-Homomorphic Encryption Scheme , in Advances in Cryptology|EUROCRYPT 2011, Lect. Notes in Comp. Sci. 6632 , (2011), Springer, 129-148.

[3] M. van Dijk, C. Gentry, S. Halevi, and V. Vaikuntanathan, Fully Homomorphic Encryption over the Integers , in Advances in Cryptology|EUROCRYPT 2010, Lect. Notes in Comp. Sci. 6110 (2010), Springer, 24-43.

[4] M. v. Dijk and A. Juels. On the impossibility of cryptography alone for privacy-preserving cloud computing. In Hot topics in Security (HotSec'10), pages 1-8. USENIX Association, 2010.

A Threshold Cryptosystem in Secure Cloud Computations

Varnovskij N.P. (ISI), Martishin S.A. (ISP RAS), Khrapchenko M.V. (ISP RAS), Shokurov A.V. (ISP RAS)

Abstract. Information security in cloud computing technology is actively investigated by the world scientific community. They uses the internet and the central remote servers to provide and maintain data as well as applications. This users' data files can be accessed and manipulated by any other users. So the problem of secure data storage and computation is actual. The modern studies in this field shows that the indicated problem is much more complex than any of the other information security problems, which are solved by well-known cryptographic methods. So, for example M. van Dijk and A. Juels in the paper "On the impossibility of cryptography alone for privacy-preserving cloud computing" described a mathematical model of the organization of cloud computing and proved that in the case of two users information protection is impossible. This result refutes the well-established point of view that the recently proposed by C. Gentry construction for fully homomorphic encryption solves at least theoretically, all the problems of information security in cloud computing. We offer an alternative model of cloud computing, in which the specified negative result does not holds. It differs from the above in the point that each subject interested in privacy, creates his own crypto server. From the point of view of users these cryptoservers are the part of the cloud. The methods of information protection, using threshold cryptosystem in this new model are investigated.

Keywords: Secure Cloud Computing, Cryptographic Protocols, fully homomorphic encryption

References

[1] C. Gentry, Fully homomorphic encryption using ideal lattices , in Proceedings of the 41st ACM Symposium on Theory of Computing|STOC 2009, ACM, New York (2009), 169-178.

[2] C. Gentry and S. Halevi, Implementing Gentry's Fully-Homomorphic Encryption Scheme , in Advances in Cryptology|EUROCRYPT 2011, Lect. Notes in Comp. Sci. 6632 , (2011), Springer, 129-148.

[3] M. van Dijk, C. Gentry, S. Halevi, and V. Vaikuntanathan, Fully Homomorphic Encryption over the Integers , in Advances in Cryptology|EUROCRYPT 2010, Lect. Notes in Comp. Sci. 6110 (2010), Springer, 24-43.

[4] M. v. Dijk and A. Juels. On the impossibility of cryptography alone for privacy-preserving cloud computing. In Hot topics in Security (HotSec'10), pages 1-8. USENIX Association, 2010.